基于 recurrent neural networks 的网约车供需预测方法 *

安　磊[a, b]，赵书良[a, b]，武永亮[a, b]，陈润资[a, b]，李佳星[a, b]

(河北师范大学 a. 数学与信息科学学院; b. 河北省计算数学与应用重点实验室，石家庄 050024)

摘　要： 以网约车订单等真实数据为数据源，结合 TensorFlow 深度学习框架，利用循环神经网络（recurrent neural networks）方法，预测网约车在未来某时间某地点的订单需求量。提出改进 LSTM RNN（长短时记忆循环神经网络）模型，经过对其优化和训练，能够有效预测网约车未来某时间某地点的供需量。对数据源进行可视化分析，排除不相关数据源干扰，以此为基础设计仿真实验。仿真实验表明，该模型的正确率比反向传播神经网络（BPNN）、回归决策树（DTR）、非线性回归支持向量机（SVR）以及随机漫步（RW）等模型高，同时，对长短间隔不同的历史数据有较好的记忆能力，在测试数据上有较强的泛化能力。

关键词： 长短时记忆循环神经网络；网约车数据；交通优化调度；TensorFlow；深度学习

中图分类号： TP391

Prediction method of supply and demand for online car based on recurrent neural networks

An Lei[a, b], Zhao Shuliang[a, b], Wu Yongliang[a, b], Chen Runzi[a, b], Li Jiaxing[a, b]

(*a. College of Mathematic & Information Science, b. Hebei Key Laboratory of Computational Mathematics & Applications Hebei Normal University, Shijiazhuang* 050024, *China*)

Abstract: Orders from online car as data sources, using TensorFlow and recurrent neural networks, to predict the supply and demand for online car at a certain point in the future. This paper presents the model of LSTM RNN, which is optimized and trained to effectively predict the supply and demand of the online car at a certain point in the future. Visual analysis of data source, help excluding uncorrelated data source, which is the basic to design simulation experiment. Simulation experiments show that the accuracy of the model proposed is higher than back propagation neural network (BPNN) and decision tree regression(DTR) , nonlinear support vector regression machine (SVR) and random walk (RW) , at the same time, the excellent memory capability of different length of historical data, and the excellent generalization capability on the test set.

Key Words: long short-term memory recurrent neural networks; online car; traffic optimization; TensorFlow; deep learning

0　引言

深度学习在人工智能研究领域的贡献逐渐增大，特别是在语音识别[1]和图像识别[2]方面，此外，研究者利用深度学习方法在其他领域的研究也得到了良好效果，例如围棋人工智能 AlphaGo[3]。深度学习算法大致上可以分为四类：深层神经网络、卷积神经网络、循环神经网络和增强学习[4]。伴随这些深度学习方法的快速发展，越来越多的深度学习框架受到人们关注，例如 TensorFlow[5]、Caffe[6]、Keras、CNTK、MXNet 等等。TensorFlow 是谷歌于 2015 年 11 月 9 日正式开源的计算框架，其计算模型能够有效地支持包含深层神经网络在内的深度学习算法，而且系统的稳定性较高[7]。TensorFlow 受到学术界和工业界的关注，例如 Geoffrey Hinton 等人利用 TensorFlow 深度学习框架进行胶囊网络的实验，结合提出的动态路由方法，对比标准卷积神经网络，在识别重叠数字问题上，取得良好的实验效果[8]；Wongsuphasawat 等人提出基于 TensorFlow 的数据流图可视化方法[9]；优步（Uber）基于 TensorFlow 的 AlexNet 深度学习模型实现无人驾驶技术[10]，Twitter、京东、小米等公司也在使用 TensorFlow。综上所述，本文使用 TensorFlow 深度学习框架进行实验。

网约车的供需变动与订单量的变化有关，同时与天气、区域配套、交通状况等因素有关，因此，供需变动具有高度非线

基金项目： 国家自然科学基金资助项目（71271067）；国家社科基金重大项目（13&ZD091）；河北省高等学校科学技术研究项目（QN2014196）；河北师范大学硕士基金资助项目（CXZZSS2017048）

作者简介： 安磊（1991-），男，河北邢台人，硕士研究生，主要研究方向为数据挖掘、智能信息处理；赵书良（1967-），男（通信作者），教授，博导，博士，主要研究方向为数据挖掘、智能信息处理（zhaoshuliang@sina.com）；武永亮（1986-），男，博士研究生，主要研究方向为数据挖掘、智能信息处理；陈润资（1981-），男，博士研究生，主要研究方向为数据挖掘、智能信息处理；李佳星（1992-），女，硕士研究生，主要研究方向为数据挖掘、智能信息处理.

性和高度随机性，准确预测供需缺口具有一定难度。人工神经网络（ANN）具有解决非线性和随机性问题的能力[11]；数据源中的数据是时间序列数据，循环神经网络（RNN）能对时间序列进行预测[12]，下文详细分析数据源的时间依赖性；在测试集中，拟预测的时间片的前半小时的数据信息已知，但是下一时间片的供需变动除与短时间内的数据有关，同时与距离该时间片较远的时间段的数据有关，因此，提出使用长短时记忆循环神经网络（LSTM RNN）作为解决供需预测问题的方法。

交通优化调度是智慧城市建设过程中的重要环节。传统的针对交通流的研究集中在两方面，一方面利用出租车 GPS 位置、车速等信息反映道路拥堵程度，另一方面利用出租车积攒的历史数据或者道路监测站收集的数据，通过参数方法或者非参数方法进行挖掘。

数据源分别从空间角度和时间角度分析，空间上，把一个城市划分为 n 个互不重合的正方形区域，并表示为集合 $D=\{d_i|d_i\in[1,58]\}$，时间上，把一天的 24 小时划分为 144 个 10 分钟长度的时间片，并表示为集合 $T=\{t_i|t_i\in[1,144]\}$，并与 $time=\{0000,0010...2350\}$ 逐一对应，有学者采用 15 分钟作为时间段[13]。基于空间和时间这两个维度，对于区域 $d_i\in D$，在时间片 $t_j\in T$，定义供需缺口 gap_{ij}，表示没有司机接单的订单数量和。

供需预测是交通流量预测的一部分，对短时交通流量的预测方法有参数化方法和非参数化方法两大类。在早期的研究中，卡尔曼过滤模型[14]作为典型的参数化方法经常被应用到交通流量的预测，参数化方法的模型结构基于某种理论假设，模型的参数预先根据经验数据计算得出。应用最广泛的参数化方法是一种叫做自回归整型滑动平均模型（ARIMA）的方法，该模型假设交通状态是静止的。ARIMA 方法又写做 ARIMA(p,d,q)，p、d、q 代表三个参数，Levin 和 Tsao 利用该方法预测高速公路上的交通流，得出 ARIMA(0,1,1)是最有效模型的结论[15]。交通流量具有非线性和随机性的特点，因此，传统的模型不能准确、有效进行预测。在人工智能领域，例如 SVM 或 SVR 等方法对无规律数据捕捉特征能力更强[16]，近年来，非参数化的方法受到重视。支持向量机（SVM）的本质是把数据通过非线性关系映射到高纬空间，然后在这个高纬空间进行线性回归，Castro-Neto 等人利用 OL-SVR 方法对交通流量在传统条件和非传统条件（例如节假日和交通事故）下进行了预测[17]；ANN 也是人工智能领域在交通流量预测问题上应用广泛的方法之一，ANN 能够解决具有高纬数据、复杂的模型结构特点的问题，并且具有较强的泛化能力和学习能力[18]。Vlahogianni 等人通过遗传算法优化了神经网络并把模型应用到了短时交通流量预测的研究中[19]；Yu 等人利用 BP 神经网络监测交通堵塞[20]；Chai 等人利用小波分析方法和神经网络预测短时交通流量[21]；Chen 利用 RBF 神经网络预测交通流量，并提出三种算法优化 RBF 权重和阈值[22]；Wang 等人利用 BP 神经网络预测公交流量并提出优化策略[23]；Yu 等人基于 RBF 神经网络提出改进人工蜜蜂群体算法用于交通预测[24]；Wang 等人提出 DeepSD 深层神经网络模型预测网约车供需量[25]。然而，现有的这些非参数化方法，均要求预先定义训练数据长度且不能改变，而且较少考虑天气、交通拥堵状况、区域配套等信息，数据源来自监测站或智能交通系统（ITS），本文使用网约车数据进行研究。

本文围绕上述问题展开研究工作，为了提高供需预测的准确率，提出一种被称为长短时记忆循环神经网络的模型(LSTM RNN)，该模型可以更有效地捕捉到数据源的非线性和随机性，并且通过记忆块克服了误差反向传播的衰变问题，同时，满足了数据源对时间序列的依赖性。而且，LSTM RNN 在测试集上取得了更高的预测精度。

1　长短时记忆循环神经网络（LSTM RNN）

循环神经网络的主要用途是处理和预测序列数据，并能够利用历史信息帮助解决当前问题，因此能够利用传统网络结构不能捕捉的信息，在数据源中，影响某一时间片内供需缺口大小的因素，除了订单需求和供给之外，还可能有交通、天气以及星期等等。

LSTM 由 Hochreiter 和 Schmidhuber[26]在 1997 年提出，循环结构的神经网络使得输入的时间序列数据能够被记忆，通过状态向量 $state$ 向后传递历史信息，因此，状态向量的定义是循环神经网络（RNN）的关键之一。然而，随着循环的进行，较早时刻的信息对当前时刻的供需缺口影响就会消失，即梯度消失问题[27]，在本文第四部分对数据源可视化的分析中，时间序列数据的长度是 $DataLenth_{d_i}=144*23=3312, d_i\in D$，影响下一时间片供需缺口大小的历史信息，与当前时刻的距离有长有短，包含距离当前时刻较近的前几个时间片的信息，也包含前几天该时刻、该区域以及周围时间片的数据信息。

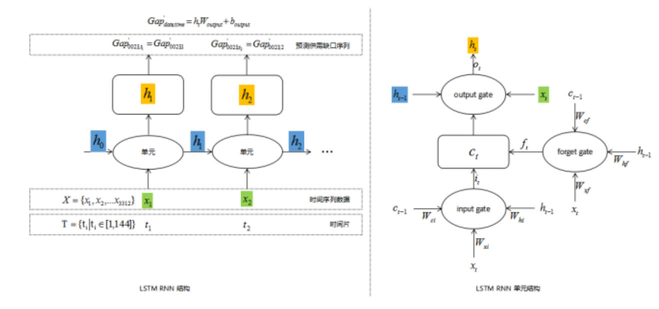


图 1　LSTM RNN 结构与单元结构示意

LSTM 结构包括 1 个输入层、1 个循环体结构和 1 个输出层，如图 1 所示。循环体结构包含 3 个门：遗忘门、输入门、输出门，门指的是以 sigmoid 为激活函数的神经网络和一个按位做乘法的操作集合，以 sigmoid 为激活函数的神经网络输出一个 0 到 1 之间的数值，描述当前输入是否可以通过这个结构。门结构可以缓解梯度消失问题，例如在第四部分的实验中，时

间序列数据（3456 维）在当前时刻输入时，若输入门关闭（sigmoid 神经网络输出层输出 0），则当前时刻的输入不会影响当前时刻的状态 $state$。

定义时间序列数据集合 $X=\{x_1,x_2,...x_{3312}\}$，状态向量 $state$ 集合 $h=\{h_1,h_2,...h_i\},i\in Z$，已知的供需缺口值集合（样本标签）是 $Y=\{y_{date,time}\}$，其中，日期 date 的范围是：$date\in\{0223,0224,...0317\}$，时间 time 的范围是：$time\in\{0000,0010,...2350\}$。

前向传播结果 $y^{'}$ 通过以下两个方程获得：

$$h_t=H(W_{input}x_t+W_{hidden}h_{t-1}+b_{hidden}) \tag{1}$$

$$Gap^{'}_{date,time}=h_tW_{output}+b_{output} \tag{2}$$

其中，$h_0=(0,0...0)$，$x_t=Gap_{date,time,d_i}$，H 函数是含 3 个门的隐藏层循环结构体，它的实现方程如下：

$$f_t=sigmoid(W_{cf}c_{t-1}+W_{hf}h_{t-1}+W_{xf}x_t+b_f) \tag{3}$$

$$i_t=sigmoid(W_{ci}c_{t-1}+W_{hi}h_{t-1}+W_{xi}x_t+b_i) \tag{4}$$

$$c_t=f_tc_{t-1}+sigmoid^{(1)}(W_{xc}x_t+W_{hc}h_{t-1}+b_c)i_t \tag{5}$$

f_t 函数是遗忘门（forget）函数，i_t 函数是输入门（input）函数，c_t 是单元激活值（cell）函数，它受到遗忘门和输入门的影响，$sigmoid^{(1)}\in[-2,2]$ 是 $sigmoid\in[0,1]$ 函数的转换函数，$sigmoid$ 函数是激活函数,它的定义如下：

$$sigmoid(x)=\frac{1}{1+e^{-x}} \tag{6}$$

输出门（output）定义如下：

$$o_t=sigmoid(W_{co}c_t+W_{ho}h_{t-1}+W_{xo}x_t+b_o) \tag{7}$$

状态向量 $state$ 受到输出门的影响，其函数如下：

$$h_t=o_tsigmoid^{(2)}(c_t),sigmoid^{(2)}\in[-1,1] \tag{8}$$

数据源的训练集在第四部分作出详细介绍，利用训练集数据对上述 LSTM RNN 模型进行训练，误差反向传播，使损失函数最小，损失函数的定义如下：

$$MAE=\frac{1}{n}\sum_{d_i}(\frac{1}{q}\sum_{t_i}\left|Gap_{date,t_i}-Gap^{'}_{date,t_i}\right|) \tag{9}$$

给定某区域 d_i 某天 $date$ 在时间片 $t_j,t_{j-1}\cdots$ 的各项数据，预测 $Gap_{date,j+1}$，并且对于 $\forall d_i\in D$。对 n 个区域和 q 个时间片，区域 $d_i\in D$ 在时间片 $t_j\in T$ 的供需缺口为 Gap_{date,t_j}，预测值为 $Gap^{'}_{date,t_j}$，以 MAE 作为损失函数。

2　LSTM RNN 算法

时间序列数据根据截断长度划分为样本数据，即拟预测时间片前阶段长度个时间片内的供需值作为样本，拟预测时间片的供需值作为标签，组成为完整的单个样本，如算法 1 所示。

算法 1 样本集构造算法

输入： 时间序列数据 seq，截断长度 TIMESTEPS。

输出： 样本集合 X,样本集合标签 Y。

```
1.def generate_data(seq,TIMESTEPS):
2.for i in range(len(seq) - TIMESTEPS):
3.X.append([seq[i:i+TIMESTEPS]])
4.Y.append([seq[i+TIMESTEPS]])
5.return X, Y
```

为了使 LSTM RNN 模型更加健壮，避免模型对训练数据的过分拟合，提高模型在测试数据上的精度，即提高模型的泛化能力，在单元激活值函数（5）的基础上，增加以概率 p 获取整数 1、以概率 $1-p$ 获取整数 0 的二分类函数[28]，目的是控制单元的激活值是否生效，根据函数的取值进行有效性判断，若 $Binary(p)=1$，则激活值有效，若 $Binary(p)=0$，则激活值失效，如下：

$$c_t=Binary(p)(f_tc_{t-1}+sigmoid^{(1)}(W_{xc}x_t+W_{hc}h_{t-1}+b_c)i_t) \tag{10}$$

构造 LSTM RNN 单层的网络结构，如算法 2（2）所示，构造二分类函数用于对单层结构的优化，如算法 2（3）所示，把 NUM_LAYERS 个经过二分类函数优化后的单层网络结构拼接成完整的 LSTM RNN 结构，如算法 2（4）所示，然后计算得出通过 LSTM RNN 前向传播得到输出 Output，经过一个全连接层网络结构,得到预测值和损失函数值，如算法 2（5）和（6）所示:

算法 2LSTM RNN 构造算法

输入： 样本集合 X，样本集合标签 Y，隐藏节点的个数 HIDDEN_SIZE，隐藏层数 NUM_LAYERS。

输出： 预测结果 prediction，损失函数 loss

```
1.def lstm_model(X,Y,HIDDEN_SIZE,NUM_LAYERS):
2.  lstm_cell←BasicLSTMCell(HIDDEN_SIZE)
3.  drop_lstm←Dropout(lstm_cell,output_keep_prob)
4.  cell←MultiRNNCell([drop_lstm] * NUM_LAYERS)
5.  Output←rnn(cell, X)
6.  prediction, loss←regression(output, Y)
7.  return prediction, loss
```

算法 1 的时间复杂度为 O(1)，算法 2 的时间复杂度包含 4 个部分，表示如下：

$$O(KH+KCS+HI+CSI)=O(W) \tag{11}$$

其中：K 表示输出单元数，H 表示隐藏单元数，C 表示记忆单元数，S 表示记忆单元大小，I 表示前馈连接记忆单元、门单

元以及隐藏单元数。LSTM RNN 的时间复杂度不依赖于网络结构和输入时间序列长度[29]，其时间复杂度为 O(W)，其中 W 表示权重数量，因此，LSTM RNN 具有高效性。

所有图灵机都可由建立在用 sigmoid 激活函数的神经元上的完全连接循环网络模拟[30]，图灵机能够计算任意可计算函数。激活函数式（6）为sigmoid函数，因此，本文提出的计算模型，理论上可以模拟供需缺口预测函数。

3　仿真实验

3.1　数据源可视化展示与分析

对数据源进行可视化分析，目的是合理划分训练集和测试集，基于此，数据源分类训练模型，通过实验对比分类前后的模型预测精度，未分类时，LSTM RNN 在测试集上的 RMSE 值为 58.171，大于分类后训练模型在测试集上的误差 43.189，发现分类之后的数据源在模型上的训练效果更好，在测试集上的误差更小，预测精度更高。数据源来自第一届 Di-Tech 算法大赛，所有数据均为真实数据。训练集 1.46GB，包括某市 2016 年连续 24 天的数据信息，本文选取部分数据用于实验。其中，订单信息表、天气信息表和 POI 信息表为数据库中直接的表信息，而区域定义表、拥堵信息表是由数据库中其他表衍生的信息。订单信息表包含字段有订单 ID（order_ID）、价格（price）、司机 ID（order_id）、用户 ID（passenger_ID）、出发地区域哈希值（start_district_hash）、订单时间戳（time）、目的地区域哈希值（dest_district_hash），根据时间间隔 10 分钟，把一天 24 小时分割成 144 个时间片，在每个时间片内，比较订单信息表的订单时间戳（time）字段划分订单所属的时间片。

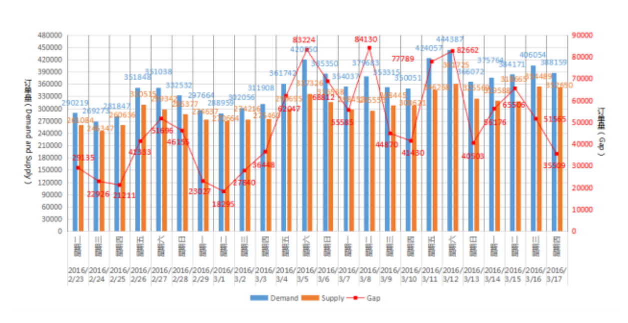


图 2　订单量日变化

定义日期 $date = \{0223,0224...0317\}$，定义需求 $Demand_{date}$，供给 $Supply_{date}$，供需缺口 Gap_{date}。在图 2 中，$Demand_{0312} = 444387$ 表示 2016 年 3 月 12 日，某市的订单需求量是 444387，相较于前后几天达到最大值，同时，供需缺口 $Gap_{0312} = 82662$ 也达到极大值，而供需缺口的最大值出现在 2016 年 3 月 8 日星期二，即 $Gap_{0308} = 84130$。

定义订单量 $ordersCount_{ij}$，其中 i 表示出发地区编号，即 $i = start_district_id \in [1,58]$，j 表示目的地区编号，即 $j = dest_district_id \in [1,58]$。在图 3 中,出发地区为 44 的集合 $\varphi_{i=44} = \{ordersCount_{44,j} \mid j \in [1,58]\}$ 与目的地区为 44 的集合 $\varphi_{j=44} = \{ordersCount_{i,44} \mid i \in [1,58]\}$ 中的元素值普遍较大，表示编号 44 的区域订单需求量大，其中 $ordersCount_{44,44} = 12554$ 表示出发地区编号（start_district_id）为 44，目的地区编号（dest_district_id）为 44 的订单量是 12554，某市当天，该区域内部往返订单量最大，在 POI 信息表中，44 号区域设施类目最多，设施数量最多；集合 $\varphi_{i\in[1,44], j\in[1,44]}$ 占据大部分总订单量，特别是集合 $\varphi_{i\in[1,22], j\in[1,22]}$ 订单量集中，其余区域订单量较小，特别是一些区域订单量为 0；集合 $\varphi_{i=j\in[1,58]}$ 中的元素值普遍较大，表示同一区域内部来往类型的订单需求量大。

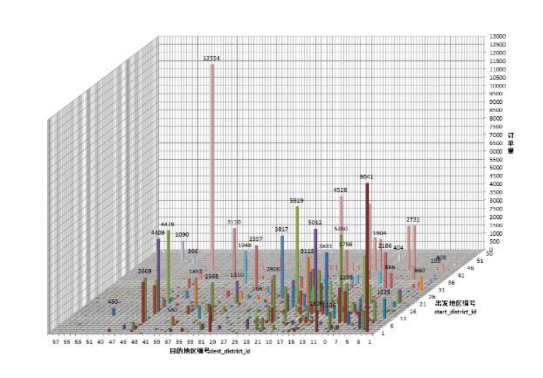


图 3　2016-02-23 订单区域分布

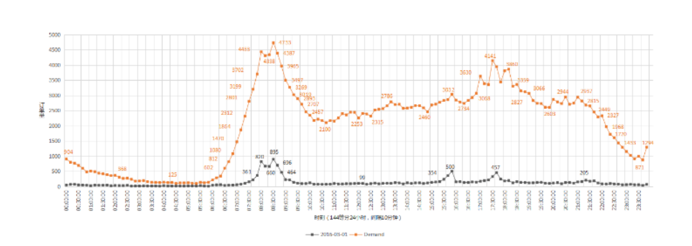


图 4　需求和供需缺口变化

定义时刻 $time = \{0000,0010...2350\}$，需求 $Demand_{time}$，供需缺口 Gap_{time}。在图 4 中，$Demand_{0830} = 4733$,表示在 2016 年 3 月 1 日 08:30-08:40，这 10 分钟时间内的订单需求达到最大值 4773，$Gap_{0830} = 895$ 表示该时间段内，供需缺口为 895，也达到了最大值，另一个需求极大值出现在 17:30，即 $Demand_{1730} = 4141$；特别地，$Demand_{0000} = 904$，$Gap_{0000} = 47$，零点之后需求量减少，供需缺口保持稳定偏低，考察同时段拥堵信息表，拥堵路段少，此外，$Demand_{2350} = 1294$ 较上一时刻有一定回升；集合 $G_1 = \{Gap_{time} \mid time \in [0720,0930]\}$ 表示 7:20-9:30 时间段内的供

需缺口值集合，集合 $G_2=\{Gap_{time}|time\in[1520,1610]\}$ 表示 15:20-16:10 时间段内的供需缺口值集合，集合 $G_3=\{Gap_{time}|time\in[1720,1820]\}$ 表示 17:20-18:20 时间段内的供需缺口值集合，$G_4=\{Gap_{time}|time\in[2100,2200]\}$ 表示 21:00-22:00 时间段内供需缺口值集合，这 4 个集合中的缺口值普遍较大，表明这 4 个时间段网约车供给不足。

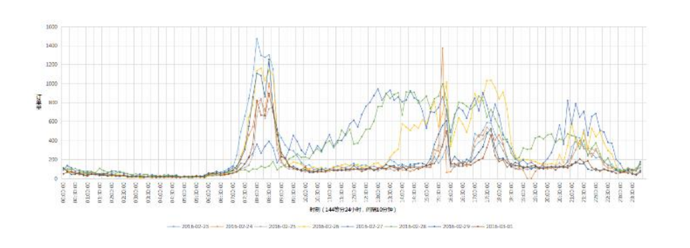


图 5　连续 8 天 Gap 变化趋势

在图 5 中，$Gap_{0223,0800}=1472$，表明 2016 年 2 月 23 日-3 月 1 日 8 天时间里，2 月 23 日 08:00-08:10 出现最大供需缺口，当天是星期二，同样出现在峰值附近网约车订单需求量增加的还有周一到周五，2 月 27 日和 2 月 28 日周末两天里，并未出现峰值，特别是星期日的 2 月 28 日，供需缺口较小；在周末这两天，网约车供需缺口在较长时间里保持增长和稳定，17:20 附近下降，在 18:40，缺口增长；特别注意在 2 月 24 日 15:30 供需缺口由 287 在 10 分钟时间内增加至 1372。订单供需缺口的变化以天为周期，周末与工作日有明显不同。通过预测未来某时间在某区域的供需缺口，可以提前采取增派措施，优化调度，缓解出行压力。

2016 年 2 月 23 日，58 个区域在 144 个时间片里的供需缺口值展示如图 6 所示，区域 44 在一天当中的几个时间区内供需缺口值较大，$Gap_{0223,1650}=244$ 表示 16:50 区域 44 的供需缺口为 244；不同区域一天内的供需变化不同，因此，训练模型时，区域信息是重要因素之一，不同区域在连续几天的供需变化存在一定规律，时间序列数据选择同一区域的连续几天的供需值。

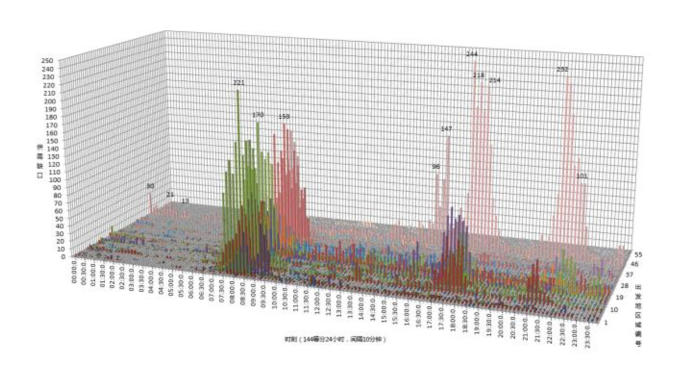


图 6　2016-02-23 时间片区域 Gap 分布

3.2　实验设计

选择数据源中前 23 天的数据信息作为训练集合，最后 1 天的数据信息作为测试集合。训练集合是前 23 天的数据，根据要预测的时间，决定是否排除周末的数据。为了与其他方法作对比，本文使用其他 4 种方法对数据分别进行了测试，4 中方法分别是：BP 神经网络方法、非线性回归支持向量机（SVR）方法、回归决策树（DTR）方法以及 Random Walk 方法。BP 神经网络是全连接神经网络，是人工神经网络最基本的方法，与 LSTM RNN 同属深度学习方法，通过二者的对比，能够比较出网络结构对模型精度的影响表现；支持向量机在分类和回归两方面都有较好的能力，核函数不同，预测性能不同，核函数包括线性核函数、多项式核函数、高斯核函数以及 sigmoid 核函数；回归决策树是一个贪心算法，即在特征空间上执行递归的二元分割；Random Walk 方法比较简单，根据当前的供需缺口值去预测下一个时间片的缺口值。

本文使用 Python 2.7、TensorFlow 0.11.0、Protobuf 3.4.0、Scikit-learn 0.19.0 以及 Scipy 0.17.0，实现 LSTM RNN 模型。为了比较 LSTM RNN 比其他 4 种方法在数据源上的预测能力更强，对比其他不同方法的均方根误差（RMSE），均方根误差可以比较真实值和预测值之间的离散程度[31]，均方根误差定义如下：

$$RMSE(Gap,Gap^{'})=[\frac{1}{n}\sum_{i=1}^{n}(Gap_i-Gap_i^{'})^2]^{\frac{1}{2}} \tag{12}$$

算法 3 计算 RMSE 算法

输入：训练轮数 TRAINING_STEPS，训练集合 seq_train，测试集合 seq_test

输出：均方根误差 rmse。

1.train_X, train_Y←
 generate_data(seq_train,TIMESTEPS)
2.test_X, test_Y←
 generate_data(seq_test,TIMESTEPS)
3.TrainedModel←
 train(train_X,train_Y,TRAINING_STEPS)
4.predicted←TrainedModel.predict(test_X)
5.rmse←sqrt((predicted - test_Y) ** 2).mean()

在训练集合中选择时间序列数据，经过算法 1 样本集构造算法，得到训练样本集合和训练样本标签集合，同理，可以获得测试样本集合和测试样本标签集合，如算法 3（1）和（2）所示，然后，经过 TRAINING_STEPS 次数迭代的训练过程，如算法 3（3）所示，获得经过训练之后的模型，使用该模型对测试数据进行预测，如算法 3（4）所示，最后，计算预测值和真实值的均方误差 RMSE。

3.3　预测精度

图 7 展示了 2016 年 3 月 17 日 44 区域内 24 小时的实际供需值和预测供需值变化，可以直观地观察到，预测值曲线较好的拟合了真实值曲线。因为 3 月 17 日是星期四，在训练集中

排除了周末的 6 天数据。模型的具体参数与均方根误差（RMSE）对比如表 1 所示。需要注意的是，表中所列的 RMSE 值，是每种方法在同等条件下，分别进行 10 次实验得到的最小值。其他四种方法的预测结果分别在图 8 展示。

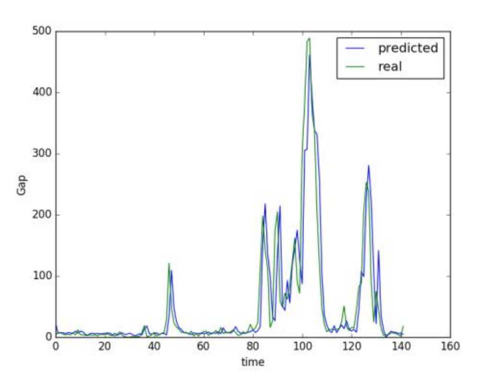


图 7　LSTM RNN 在测试集合预测值与真实值曲线

表 1　LSTM RNN 与四种方法对比

模型名称	RMSE	描述
LSTM RNN	43.189	输入节点数 1,隐藏层数 2,隐藏层节点数 $h=68$,截断长度为 2
BPNN	61.218	输入节点数 5,隐藏层数 1,隐藏层节点数 50
SVR	92.217	高斯核函数,特征向量长度为 1
DTR	98.992	特征向量长度为 2
RW	256.393	2 个随机方向,随机范围为(-200,200),值范围为(0,600)

通过对比其他四种预测方法，BPNN 方法的均方根误差值为 61.218，SVR 方法的均方根误差值为 92.217，DTR 方法的均方根误差值为 98.992，RW 方法的均方根误差值为 256.393。其中，LSTM RNN 方法的误差在 10 次实验中，最大值是 46.377，表明 LSTM RNN 在测试集合上的误差变化小，预测性能比较稳定；BPNN 方法的误差在 10 次实验中的最大值为 95.221，误差的波动幅度较大，表明 BPNN 方法的预测性能不稳定；SVR 方法的误差在 10 次实验中变化较小，RMSE 值稳定在 92 附近，说明 SVR 方法在测试集合上的预测稳定性较高；DTR 方法的误差在 10 次实验中，误差最大值为 99.339，误差具有一定的波动，但是比 BPNN 的误差波动小；RW 方法在测试集合上的误差变化较大，说明根据预测时间片前一个时间片对下一时刻进行预测的方法随机性较强，误差较大，而且较不稳定。

通过比较发现，LSTM RNN 的 RMSE 值最小，说明相比其他方法在数据源上预测精度最高。其他 4 种回归方法，结构不同，但在解决预测问题都是比较常用的方法，但在处理时间序列问题上，LSTM RNN 自身的结构设计使得它能够比其他回归算法具有更高的预测精度。

LSTM RNN 的结构对模型的预测能力有直接影响，特别是隐藏层的层数和每一层隐藏层节点的个数。通过设置不同的值，找到最优的 LSTM RNN 结构，在隐藏层数为 2 的基础上，观察不同节点数对 RMSE 大小的影响，如图 9，当隐藏层节点数较小时，RMSE 较大，随着隐藏层节点数越来越大，RMSE 也越来越小，当隐藏层节点数为 68 时，RMSE 达到最小，此后，隐藏层节点数越来越大，RMSE 稳定在 50 附近。

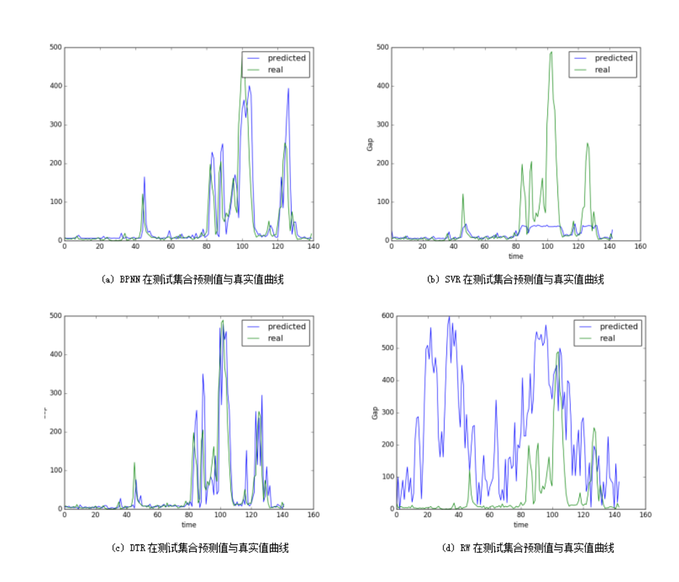


图 8　4 种方法在测试集合预测值与真实值曲线

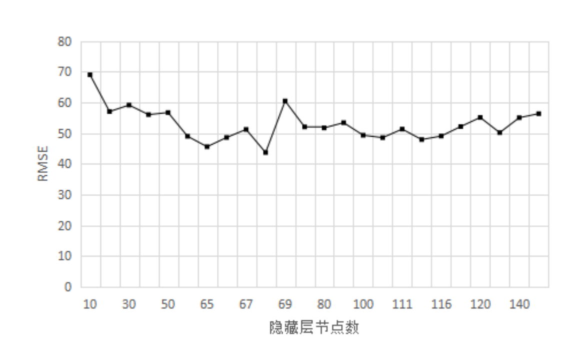


图 9　RMSE 随隐藏层节点数变化

3.4　截断长度与 RMSE 大小的关系

截断长度就是要预测时间片前的时间片的个数，比如预测时间片前 30 分钟的截断长度就是 3（时间片大小是 10min），在同等的模型结构条件下，不同的截断长度对模型的预测精度也会产生比较大的影响，如图 10 所示。

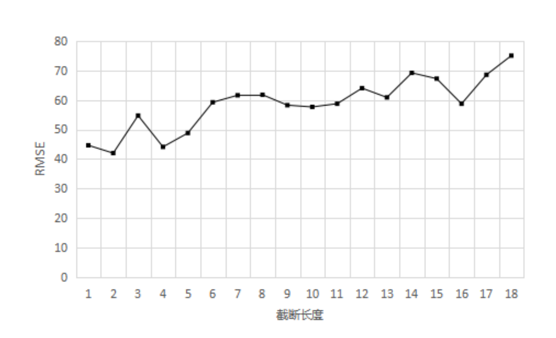


图 10　RMSE 随截断长度变化

从图中可以看出，$h=2$时，RMSE 最小为 43.189，即预

测时间片前 20 分钟的数据对预测精度的影响最大，图中 h 最长为18，表示预测时间片前3小时的数据信息。实验表明，截断长度为1时，LSTM RNN 的预测值与真实值间的均方根误差 RMSE 为44.645，截断长度为3时均方根误差为54.687，截断长度为4时均方根误差为44.112，当截断长度大于4时，均方根误差增大，并且随着截断长度逐渐增大，RMSE 的值也逐渐增大。长短时记忆循环神经网络，能够对不同长度的历史数据进行记忆，是 LSTM RNN 的优势之一。

4　结束语

本文提出了用 LSTM RNN 模型预测订单供需缺口的方法，LSTM RNN 具有处理时间序列数据的能力，隐藏层节点数量、截断长度以及隐藏层数是模型结构的重要参数，对模型在数据源上的预测精度有着重要影响。本文通过仿真实验，发现最佳的截断长度为 2 即预测时间片前 20 分钟的订单信息对预测值的影响最大。通过对训练数据的可视化分析，发现周末两天的订单信息与工作日的订单信息差别较大，因此在实验中，为了提高精度，分类处理训练数据集合。为了证明 LSTM RNN 具有更好的预测精度，通过与其他4种回归方法进行比较，分别是反向传播神经网络（BPNN）、非线性支持向量机（SVR）、回归决策树（DTR）、随机漫步（RW），比较这几种方法在测试数据集合上的均方根误差（RMSE），实验结果显示，LSTM RNN 的均方根误差最小，说明了 LSTM RNN 对订单供需预测具有更高的精度，验证了 LSTM RNN 具有更好的泛化能力以及对长短间隔的历史数据的记忆能力。将来的主要工作集中在两个方面，一方面，在模型中加入优惠活动、特殊节日等因素对订单供需缺口的影响，并考虑实际情况，在损失函数的设计中加入偏置；另一方面，改进模型结构，考虑实时的供需变动因素。

参考文献：

[1] Sainath T N, Weiss R J, Wilson K W. Multichannel signal processing with deep neural networks for automatic speech recognition [J]. IEEE//ACM Trans on Audio, Speech, and Language Processing, 2017, 25 (5): 965-979.

[2] Goodfellow I, Bengio Y, Courville A, Deep learning (adaptive computation and machine learning series) [M]. Canbridge: MIT Press, 2016.

[3] Silver D, Huang A, et al. , Mastering the game of go with deep neural networks and tree search [J]. Nature, 2016, 529: 484-489.

[4] Yi H, Jung H J, Bae S. Deep neural networks for traffic flow prediction [C]// Proc of IEEE International Conference on Big Data and Smart Computing. 2017: 328-331.

[5] Mo Y J, Kim J, Kim J K, et al. Performance of deep learning computation with TensorFlow software library in GPU-capable multi-core computing platforms [C]// Proc of the 9th International Conference on Ubiquitous and Future Networks. 2017: 240-242.

[6] Ichinose A, Oguchi M, Takefusa A, et al. Evaluation of distributed processing of caffe framework using poor performance device [C]// Proc of IEEE International Conference on Big Data. 2016: 3980-3982.

[7] Abadi M, Agarwal A, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems [R]. Preliminary White Paper, 2015.

[8] Hinton G, Sabour S, Frosst N. Dynamic routing between capsule [C]// Proc of Conference on Neural Information Processing Systems. 2017.

[9] Wongsuphasawat K, Smilkov D, et al. Visualizing dataflow graphs of deep learning models in TensorFlow [J]. IEEE Trans on Visualization and Computer Graphics, 2018: 24 (1): 1-12.

[10] Gallardo N, et al. Autonomous decision making for a driver-less car [C]// Proc of the 12th System of Systems Engineering Conference. 2017

[11] Gao Feng. Network traffic prediction based on neural network [C]// Proc of International Conference on Intelligent Transportation, Big Data and Smart City. 2015: 527-530.

[12] Chen Dawei. Research on traffic flow prediction in the big data environment based on the improved RBF neural network [J]. IEEE Trans on Industrial Informatics, 2017, 13 (4): 2000-2008.

[13] Highway Capacity Manual. Transportation research board, National Research Council, Washington, DC, 2000: 113.

[14] Karlaftis M, E. Vlahogianni. Statistical methods versus neural networks in transportation research: Differences, similarities and some insights [J]. Transportation Research Part C: Emerging Technologies, 2011, 19 (3): 387-399.

[15] Levin M, Tsao Y D. On forecasting freeway occupancies and volumes (abridgment) [J]. Transportation Research Record, 1980, 773.

[16] Yu Haiyang, Wu Zhihai, Chen Dongwei, et al. Probabilistic prediction of bus headway using relevance vector machine regression [J]. IEEE Trans on Intelligent Transportation Systems, 2017, 18 (7): 1772-1781.

[17] Castro-Neto M, Jeong Y S, Jeong M K, et al. Onlinesvr for short-term traffic flow prediction under typical and atypical traffic conditions [J]. Expert systems with applications, 2009, 6 (3): 6164-6173.

[18] Chan K Y, Dillon T, Chang E, et al. Prediction of short-term traffic variables using intelligent swarm-based neural networks [J]. IEEE Transs on Control Systems Technology, 2013, 21 (1): 263-274.

[19] Vlahogianni E I, Karlaftis M G, Golias J C. Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach [J]. Transportation Research Part C: Emerging Technologies, 2005, 13 (3): 211-234.

[20] Yu Xiaohan, Xiong Shengwu, Xiang Jianwen, et al. A campus traffic congestion detecting method based on BP neural network [C]// Proc of the 2nd International Symposium on Dependable Computing and Internet of Things. 2015.

[21] Chai Yanchong, Huang Darong, Zhao Ling. A short-term traffic flow prediction method based on wavelet analysis and neural network [C]// Proc of Chinese Control and Decision Conference. 2016: 7030-1034.

[22] Chen Dawei. Research on traffic flow prediction in the big data environment

based on the improved RBF neural network [J]. IEEE Trans on Industrial Informatics, 2017, 13 (4): 2000-2008.

[23] Wang Peng, Zhao Gang, Yao Xingren. Applying back-propagation neural network to predict bus traffic [C]// Proc of the 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery. 2016: 752-756.

[24] Yu Wanxia, Liu Lina; Zhang Weicun. [C]// Proc of the 8th International Conference on Intelligent Networks and Intelligent Systems. 2015: 141-144.

[25] Wang Dong, Cao Wei, Li Jian. DeepSD: supply-demand prediction for online car-hailing services using deep neural networks [C]// Proc of the 33rd International Conference on Data Engineering 2017: 243-254.

[26] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997, 9 (8): 1735-1780.

[27] Ma Xiaolei, Ding Chuan, Luan Sen, et al. Prioritizing influential factors for freeway incident clearance time prediction using the gradient boosting decision trees method [J]. IEEE Trans on Intelligent Transportation Systems, 2017, 18 (9): 2303-2310.

[28] Zaremba W, Sutskever I, Vinyals O. Recurrent neural network regularization [J/OL]. Eprint Arxiv, 2014. https://arxiv. org/abs/1409. 2329v5

[29] Schmidhuber J. Learning to control fast-weight memories: an alternative to dynamic recurrent networks [J]. Neural Computation, 1992: 4 (1): 131-139.

[30] Siegelmann H T, Sontag E D. Some recent results on computing with'neural nets' [C]// Proc of the 31st IEEE Conference on Decision and Control. 1992: 1476-1481.

[31] Fouladgar M, Parchami M, Elmasri R, et al. Scalable deep traffic flow neural networks for urban traffic congestion prediction [C]// Proc of International Joint Conference on Neural Networks. 2017: 2251-2258.